

Heterogeneity or True State Dependence in Poverty - The tale told by twins

by William Nilsson¹

Department of Economics, Umeå University

This study focuses on the persistence of poverty in Sweden. The purpose is to distinguish between two different reasons why poverty could persist on an individual level. By using a sample of identical twins, this study takes advantage of the similarity within pairs of twins to separate family specific heterogeneity from true state dependence, where the experience of poverty leads to a higher risk of future poverty. The results, based on a four variate probit model, show the importance of true state dependence in poverty. When using the information on whether an individual received social assistance as a measure of poverty, family specific heterogeneity explains between 24 and 31 percent of the poverty persistence in the sample.

JEL: I32, C35, D31

Keywords: poverty, heterogeneity, state dependence, twins, multivariate probit

¹I wish to thank seminar participants at the Department of Economics, Umeå University, Sweden. In particular I would like to thank Karl-Gustaf Löfgren, Kurt Brännäs and Tomas Sjögren. I also want to thank Stephen P. Jenkins for providing valuable comments on a late draft of the paper. Further, I want to thank Chris Hudson for improving the language. Financial support from the Wallander-Hedelius-Browaldh Foundation is gratefully acknowledged. All data, except for the identification of the twin sample and information of whether the twins are identical or not, came from Statistics Sweden. The twin information came from the Swedish Twin Registry. The Swedish Twin Registry is supported by grants from the Swedish Research Council. To contact the author, use william.nilsson@econ.umu.se.

1. Introduction

A state of poverty is often referred to as a situation in which there is a lack of resources to achieve a reasonable standard of living. An individual living in such situation is usually more likely to continue in this state in following years. In the work against poverty and social exclusion, it is important to know who is at risk of becoming poor, and what characteristics make poverty persistent at an individual level. In designing an efficient system against poverty, it is important to know to what extent poverty persists due to heterogeneity or how much is it due to true state dependence. In this case, true state dependence refers to a situation in which *the experience* of poverty causes a subsequently higher risk of continuing to be poor. The individual can, for example, lose motivation or develop health problems making poverty more probable in the future. Heterogeneity could also be the explanation for the persistence of poverty. In this case, it could be characteristics that are specific for the individual, such as low level of education, that increase the risk of poverty. These characteristics could have their origin in the environment in which the individual grew up. If the characteristics that are important for the risk of poverty persist, the risk of poverty will also persist.

If true state dependence is relatively more important than heterogeneity, an effective policy could focus on preventing people from becoming poor, since once there, they are likely to remain in poverty no matter what their initial characteristics were. On the other hand, if heterogeneity explains the persistence in poverty, it is important to focus on changing the characteristics that keep the individual at a high risk of being poor.

The purpose of this study is to distinguish between true state dependence and family specific heterogeneity. This is done using a data set of identical twins. The idea is to take the similarity of the twins into account as their innate abilities are the same, and the social backgrounds are very similar for pairs of identical twins. The method relies on the assumption that *the experience* of poverty for one twin does not affect the probability of poverty for the sibling twin the following year. If the assumption is valid, a higher probability of poverty depending on whether the sibling twin was poor or not, is due to the similarity between the twins. This information can be used to distinguish between family specific heterogeneity and true state dependence.

Previous studies investigating true state dependence and heterogeneity as explanations for poverty persistence have, to my knowledge, never taken advantage of information on twins or siblings. Without information for twins, Stewart & Swaffield (1999) estimate a bivariate probit model consisting of two equations. The first equation models the probability of low pay in a base year. The second equation is a transition equation and is the probability of low pay the following year, for those who were low paid in the base year. The equations are estimated simultaneously to address the potential problem that those who are low paid in the base year are not necessarily a random sample of the total population. To assume exogeneity of low pay in the base year used in the transition equation, and then not to take this into account, could lead to biased estimates. Stewart & Swaffield (1999) identify the state dependence effect by using the estimates for the coefficients in the transition equation.²

²The main idea is the same as in Cappellari & Jenkins (2004) and is explained in detail in section 2 in this study.

Cappellari & Jenkins (2004) estimate a similar model but include both the transition out of and the transition into low income. They also include a third equation that takes into account that survey data can be affected by non random attrition.³ Individuals can leave the panel or fail to answer all important questions. If some individuals, for example poor individuals in the base year, are more likely to leave the panel, the estimates could be biased, if this non-random attrition is not taken into account. Both the studies mentioned use survey data to investigate low pay or low income transitions.

The main contribution of this study is to distinguish between family specific heterogeneity and true state dependence in poverty using a method based on twins. Using information for twins is a new and innovative way to study persistence in a state. The measures presented for family specific heterogeneity and true state dependence are appealing since the sibling twin provides a reference case with very similar unobserved characteristics. Further, the study is based on administrative data instead of survey data, which considerably reduces the rate of attrition.⁴ The results reveal that even though heterogeneity plays its role, true state dependence is relatively more important for the persistence of poverty for the sample of twins used here.

The remainder of this paper is structured in the following way. Section 2 explains the model. Section 3 provides information on the data and discusses measures of poverty. The results from the empirical investigation are included in section 4. The concluding remarks can be found in section 5.

³Panel attrition refers to a case where individuals leave the panel, and accordingly cannot completely contribute to the estimates.

⁴In this study individuals can still leave the panel through migration abroad or death. These reasons are, however, not that common.

2. Model

In this section the model is described. Firstly, the main idea of the new method using twins is explained. Then the model, based on a multivariate probit model, is described. Finally, the measures for family specific heterogeneity and true state dependence are defined based on the regression model.

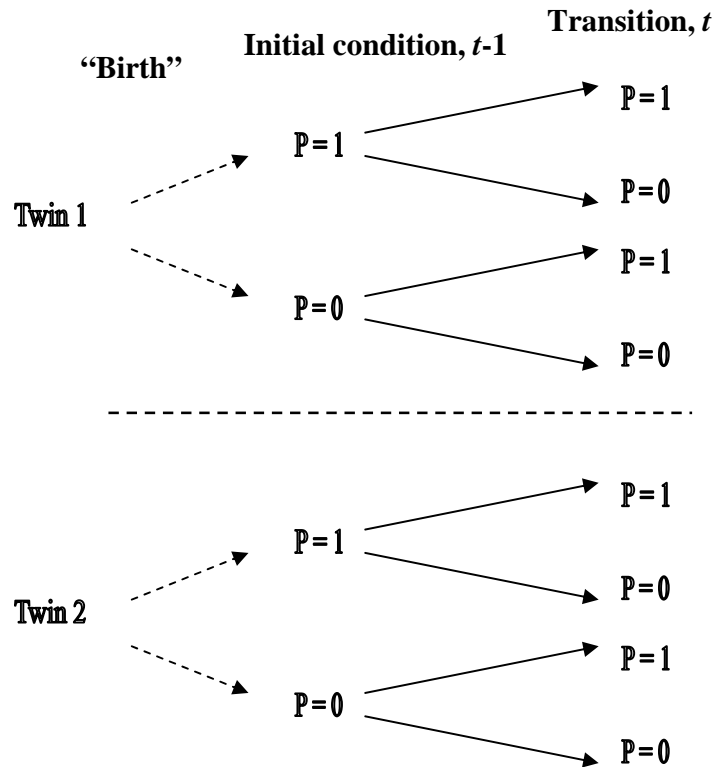


Figure 1. Overview of poverty transitions

Figure 1 gives an overview of poverty transitions for individuals. Individuals are observed as either poor ($P = 1$) or not ($P = 0$) in period

$t - 1$ and t .⁵ If individuals who are poor in $t - 1$ are compared with those who were not poor in $t - 1$, we expect that the average probability of being observed as poor in t will differ between the groups. If poverty persists on an individual level, the group that was poor in $t - 1$ will have a higher average probability of being poor in t compared with the group that was not poor in t . Individuals in the two groups are expected to differ both because of their initial characteristics and because of their experience of poverty in $t - 1$. Thus, the measures do not tell us what part of the poverty persistence is caused by heterogeneity and what part is due to true state dependence.

The idea in this study is to compare the average probability of being observed as poor depending on whether the twin sibling was observed as poor in the previous year or not. The crucial assumption for the analysis is that one twin's experience of poverty, *in itself*, does not affect the probability of the sibling twin being poor the following year. If true state dependence comes from health problems this seems reasonable. This assumption is hereafter called the Twin-State Independence Assumption. Instead of comparing two groups on the basis of their own base year poverty status, the new method compares individuals depending on the status of their twin sibling. Thus, the first group includes twins who have a twin sibling that was poor the previous year and the second group includes those whose sibling twin was not poor. If the group of twins, with a twin who was poor the previous year, have a higher average probability of being poor, this is likely to be due to the fact that the sibling twins have some important characteristics in common that increase the risk of poverty for both of them. True state dependence cannot explain a higher probability for the first group since

⁵Note that poverty here is defined as a binary variable while it is certainly true that there are degrees of poverty. The situation for the poor could, for example, be very different even though the number of poor would be the same. One reason for using a

the groups are not compared with respect to *their own* poverty status the previous year.

This method of distinguishing between true state dependence and heterogeneity gives a lower bound for the persistence of poverty caused by heterogeneity. The reason is that even identical twins are different, and these differences cannot be captured with this method. Identical twins are identical with respect to innate abilities and with respect to many acquired characteristics based on growing up in the same environment. There are, however, experiences that differ between the twins. One of these could be the experience of poverty. Since all acquired characteristics that differ between the twins, except from the experience of poverty, should be labelled heterogeneity, the method only identifies a lower bound for the heterogeneity. It is important for the understanding of the method to be aware that *family specific* heterogeneity in this study refers to differences towards the rest of the population, and not towards the sibling twin.

To distinguish between true state dependence and family specific heterogeneity, this paper extends the models for low pay/income transition used in Stewart & Swaffield (1999) and Cappellari & Jenkins (2004) to include twin sibling homogeneity as a way of revealing heterogeneity towards the rest of the population. As in the above mentioned articles, the first part of the model refers to the risk for poverty in a base period, $t - 1$. This is to deal with the potential problem that the poverty status is not exogenously determined. The latent poverty propensity, P_{sit-1}^* , is assumed to be determined by

discrete variable in research is, however, that a binary variable is often used when policies are constructed and later evaluated.

$$P_{1it-1}^* = \beta_1' \mathbf{x}_{1it-1} + \mu_{1i} + \delta_{1it-1} \quad (1)$$

$$P_{2it-1}^* = \beta_2' \mathbf{x}_{2it-1} + \mu_{2i} + \delta_{2it-1} \quad (2)$$

$$P_{sit-1} = I(P_{sit-1}^* > 0) \\ s = 1, 2, i = 1, \dots, N, \text{ and } t = 2, \dots, T$$

Subindex $s = 1$ indicates the first group of twins, and $s = 2$ indicates the second group of twins consisting of the sibling twin of $s = 1$. Subindex i refers to the pair of twins $i = 1, \dots, N$ and $t = 1, \dots, T$ refers to different periods.⁶ Individuals are only observed to be either poor, $P_{sit-1} = 1$, or not, $P_{sit-1} = 0$. $I(P_{sit-1}^* > 0)$ is an indicator function which takes the value 1 if the inequality is satisfied, and zero otherwise. Individual characteristics that are assumed to influence the poverty status are included in \mathbf{x}_{sit-1} , and β_s is a vector of the parameters to be estimated.⁷ u_{sit-1} is an error term which includes μ_{si} , an individual specific effect, and δ_{sit-1} , an orthogonal white noise error. The error term is assumed to follow a standard normal distribution, $u_{sit-1} \sim N(0, 1)$.

Given the poverty status in period $t - 1$, the next step is to model the transition equations, i.e. the transition to the next period, t , in *Figure 1*. The transition equations model the probability of remaining in poverty for those who were poor in the previous year, and the probability of entering into poverty for those who not were poor the previous year.

⁶At this point, it can be assumed that the data consist of two periods, as indicated in *Figure 1*. The case where more observations are present for each individual will be discussed briefly later.

⁷In the estimation, there is no reason to believe that β_1 should differ systematically from β_2 . It is, accordingly, possible to constrain β_1 to be equal to β_2 .

$$P_{1it}^* = [P_{1it-1}\gamma'_{11} + (1 - P_{1it-1})\gamma'_{12}]z_{1it-1} + \tau_{1i} + \zeta_{1it} \quad (3)$$

$$P_{2it}^* = [P_{2it-1}\gamma'_{21} + (1 - P_{2it-1})\gamma'_{22}]z_{2it-1} + \tau_{2i} + \zeta_{2it} \quad (4)$$

$$P_{sit} = I(P_{sit}^* > 0)$$

$$s = 1, 2, i = 1, \dots, N \text{ and } t = 2, \dots, T$$

The included parameters, i.e. those attached to the explanatory variables, z_{sit-1} , in equation (3) and (4), are allowed to vary in magnitude depending on the poverty status in the previous period.⁸ This setup follows Cappellari & Jenkins (2004) apart from the trivial extension of the notation for each twin. A further difference is that they also model a panel retention equation, i.e. whether the individual remains in the panel which is not necessary in the present study. The error term, v_{sit} consists of two parts, τ_{si} , an individual specific effect, and ζ_{sit} , an orthogonal white noise error. The error term is again assumed to be $v_{sit} \sim N(0, 1)$.

In general, many of the variables included in x_{sit-1} are also used in z_{sit-1} . The estimation could be identified without excluded variables if a nonlinear functional form is present. However, to avoid relying on the assumption of nonlinearity, it is appropriate to include instruments in the first period equation (1 and 2), which can be excluded from the equation for the transition (3 and 4). That is, variables should be

⁸As for β_s , there are no reasons to expect γ'_{s1} and γ'_{s2} to differ systematically for the different sample of twins, i.e. $s = 1, 2$. γ'_{11} can, accordingly, be constrained to equal γ'_{21} and γ'_{12} can be constrained to equal γ'_{22} .

included that affect the probability of the poverty status in the base year but not the transition, given information for the status of poverty in the base year. In Stewart & Swaffield (1999) and Cappellari & Jenkins (2004), these instruments are variables relating to the parents' socioeconomic group when the respondent was 14 years old. Similar variables are used in this study.

The joint distribution of the error terms, $u_{1it-1}, u_{2it-1}, v_{1it}, v_{2it}$, is assumed to have a correlation matrix. The correlation between the error terms for the different equations are labelled as follows,

$$Corr \begin{pmatrix} u_{1it-1}, u_{1it-1} & & & \\ u_{2it-1}, u_{1it-1} & u_{2it-1}, u_{2it-1} & & \\ v_{1it}, u_{1it-1} & v_{1it}, u_{2it-1} & v_{1it}, v_{1it} & \\ v_{2it}, u_{1it-1} & u_{2it-1}, u_{2it-1} & v_{2it}, v_{1it} & v_{2it}, v_{2it} \end{pmatrix} \equiv$$

$$\begin{pmatrix} 1 & & & \\ \rho_{21} & 1 & & \\ \rho_{31} & \rho_{32} & 1 & \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}$$

where the subindices indicate between which two equations the correlation refers. In the case where none of the error terms is correlated between the equations, it would be possible to estimate the four equations by four separate probit models. In this study the equations are, however, estimated simultaneously, and the correlation between the error terms of the equations are not constrained in any way. This gives an opportunity to actually test whether the correlations are different from zero or not.

The log-likelihood function for each pair of twins, $i = 1, \dots, N$, and $t = 1, 2$, is;

$$\log L_i = \log \Phi_4(\mu_i; \Omega) \quad (5)$$

where $\Phi_4(\mu_i; \Omega)$ is a standard four-variate normal cdf, with

$$\mu_i = \{K_{i1}\beta'_1 \mathbf{x}_{i1t-1}, K_{i2}\beta'_2 \mathbf{x}_{i2t-1}, K_{i3}[P_{i1t-1}\gamma'_{11} + (1 - P_{i1t-1})\gamma'_{12}]\mathbf{z}_{i1t-1}, \\ K_{i4}[P_{i2t-1}\gamma'_{21} + (1 - P_{i2t-1})\gamma'_{22}]\mathbf{z}_{i2t-1}\}$$

where, $K_{i1} = 2P_{i1t-1} - 1$, $K_{i2} = 2P_{i2t-1} - 1$, $K_{i3} = 2P_{i1t} - 1$, and $K_{i4} = 2P_{i2t} - 1$.

The matrix, Ω is symmetric;

$$\Omega = \begin{pmatrix} 1 & K_{i2}K_{i1}\rho_{21} & K_{i3}K_{i1}\rho_{31} & K_{i4}K_{i1}\rho_{41} \\ K_{i2}K_{i1}\rho_{21} & 1 & K_{i3}K_{i2}\rho_{32} & K_{i4}K_{i2}\rho_{42} \\ K_{i3}K_{i1}\rho_{31} & K_{i3}K_{i2}\rho_{32} & 1 & K_{i4}K_{i3}\rho_{43} \\ K_{i4}K_{i1}\rho_{41} & K_{i4}K_{i2}\rho_{42} & K_{i4}K_{i3}\rho_{43} & 1 \end{pmatrix}$$

As in Cappellari & Jenkins (2004), the multivariate standard normal distribution function is evaluated using a simulation method based on the GHK simulator.⁹ (See Cappellari & Jenkins, 2003, for a detailed description on the method.)

⁹To estimate the model, I used the Stata program `mvprobit` written by Cappellari & Jenkins (2003). Stata users can obtain the program by typing `findit mvprobit` at the Stata prompt. GHK = Geweke-Hajivassiliou-Keane.

When having observations for more than two periods it is possible to estimate a pooled model for all periods. Doing this would, however, be to use a large amount of the observations for both the initial condition and for the transition equation. Having more than two periods would, in fact, extend the tree in *Figure 1* to the right for poverty status in $t + 1$, $t + 2$, etc. As the new figure would indicate, each additional observation would have a longer history of transitions back to the initial condition. Accordingly, to pool the observations would involve basing the likelihood on the wrong set of information. Even if only a few more periods are added, taking the correct set of information into account would considerably complicate the model. If a panel of data are present, an alternative to pooling all the observations is to model two periods at a time. This would mean several different regressions and several different measures of state dependence. However, if the number of individuals in the data set is small, each regression would be based on few observations.

A second alternative is to pool observations, but to avoid using the same observations for both the initial condition equation and the transition equation. For example, in a panel of four years, $t - 1$ would be used as base year for the transition in t , and $t + 1$ would be the base year for the transition in $t + 2$. This is the method used in this study and also the method used in Cappellari & Jenkins (2004). The standard errors for the estimates are corrected for repeated observations from the same twins over the years. Before focusing on the data and estimation, it is important to know how to use the model to estimate measures of state dependence and heterogeneity.

2.1 Distinguishing true state dependence and heterogeneity

The first step towards distinguishing between true state dependence and heterogeneity in the persistence of poverty is to define transition probabilities such as those used in for example, Cappellari & Jenkins (2004). From the model, it is possible to calculate the probability of being poor at t , conditional on being poor in the previous year. This is the poverty persistence rate, i.e. to what extent individuals remain in poverty.

$$\Pr(P_{sit} = 1 | P_{sit-1} = 1) = \frac{\Phi_2(\gamma'_{s1} \mathbf{z}_{sit-1}, \beta'_s \mathbf{x}_{sit-1}; \rho)}{\Phi(\beta'_s \mathbf{x}_{sit-1})} \quad (6)$$

where $\rho = \rho_{31}$ for $s = 1$ and $\rho = \rho_{42}$ for $s = 2$. ρ_{31} refers, as explained above, to the correlation of the error terms for the equations for the first set of twins. If the poverty persistence rate is estimated for the sibling twins, the counterpart for $s = 2$, ρ_{42} , is used. Further, $\Phi(\cdot)$ and $\Phi_2(\cdot)$ are the cumulative density functions of univariate and bivariate standard normal distributions. The poverty entry rate is the probability of being poor at t , conditional on not being poor the previous year.

$$\Pr(P_{sit} = 1 | P_{sit-1} = 0) = \frac{\Phi_2(\gamma'_{s2} \mathbf{z}_{sit-1}, -\beta'_s \mathbf{x}_{sit-1}; \rho)}{\Phi(-\beta'_s \mathbf{x}_{sit-1})} \quad (7)$$

where $\rho = -\rho_{31}$ for $s = 1$ and $\rho = -\rho_{42}$ for $s = 2$. With these different measures of probabilities of poverty, it is possible to calculate measures of state dependence and heterogeneity.

Cappellari & Jenkins (2004) estimate aggregate state dependence, ASD , as

$$ASD_s = \left(\frac{\sum_{i \in (P_{sit-1}=1)} Pr(P_{sit}=1|P_{sit-1}=1)}{\sum_i P_{sit-1}} \right) - \left(\frac{\sum_{i \in (P_{sit-1}=0)} Pr(P_{sit}=1|P_{sit-1}=0)}{\sum_i (1-P_{sit-1})} \right)$$

ASD is the difference in the average probability of being poor for those who were poor in the previous year and the average probability of being poor for those who were not poor the previous year.¹⁰ ASD measures state dependence without taking into consideration that the poor and the non poor the previous year could be very different.

Another measure that Cappellari & Jenkins (2004) use is what they call genuine state dependence, GSD . GSD is an average, over all individuals, of the difference between the predicted probability of being poor conditional on being poor, and the predicted probability of being poor conditional on not being poor, the previous period.

$$GSD_s = (1/N) \sum_{i=1}^N Pr(P_{sit} = 1|P_{sit-1} = 1) - Pr(P_{sit} = 1|P_{sit-1} = 0)$$

¹⁰Note that $\sum_i P_{sit-1}$ is just the number of individuals poor the previous year, since $P_{sit-1} = 1$ for those, and $P_{sit-1} = 0$ for the non poor. In the same way $\sum_i (1-P_{sit-1})$ is the number of individuals that were not poor. Thus, each term in ASD_s is just an average for each subgroup.

This measure controls for heterogeneity since individual specific probabilities are averaged. The individual specific probabilities differ due to differences between the estimated parameters (γ'_{s1} and γ'_{s2} in equation (3) and (4)) depending on the poverty status the previous year. Cappellari & Jenkins (2004) state that the measure controls for both observed and unobserved heterogeneity. While it is easy to see how the observed heterogeneity, i.e. for those explanatory variables included, is taken into account, it is not obvious how the inclusion of ρ_{31} (or ρ_{42}) in the individual specific probabilities takes care of all unobserved heterogeneity. Using information for twin siblings, an intuitively more appealing measure can be calculated.

Equations for twin siblings enable probability expressions to be devised that can be used to construct a measure to identify the part of poverty persistence that is due to family specific heterogeneity. The probability of the first twin being poor at t , conditional on the second twin being poor the previous year, can be calculated as

$$\Pr(P_{1it} = 1 | P_{2it-1} = 1) = \frac{\Phi_2(\gamma'_1 \mathbf{z}_{1it-1}, \beta'_2 \mathbf{x}_{2it-1}; \rho_{32})}{\Phi(\beta'_2 \mathbf{x}_{2it-1})} \quad (8)$$

where γ'_1 refers to γ'_{11} and γ'_{12} estimated in equation (3).¹¹ The appropriate γ' is used depending on whether the first twin was poor or not the previous year. The probability of the first twin being poor at t , conditional on the second twin not being poor the previous year is calculated as

¹¹Note that the probability of the *second* twin being poor at t , conditional on the *first* twin being poor the previous year is calculated with ρ_{41} . The changes in variables and coefficients are obvious.

$$\Pr(P_{1it} = 1 | P_{2it-1} = 0) = \frac{\Phi_2(\gamma_1' \mathbf{z}_{1it-1}, -\beta_2' \mathbf{x}_{2it-1}; -\rho_{32})}{\Phi(-\beta_2' \mathbf{x}_{2it-1})} \quad (9)$$

If the probability to be poor is higher for the first twin conditional on whether the second twin was poor the previous year, it can only be due to homogeneity within the pairs of twins as long as the Twin-State Independence Assumption is valid. In other words, the assumption that the experience of poverty for one twin does not *in itself* affect the probability of poverty the following year for the twin sibling. The twin pair has to have common characteristics or experiences that increase the risk of poverty. State dependence cannot be a reason why the probability would be higher, as it was the twin sibling that experienced poverty the previous year.

For the first sample of twins the new measure is calculated according to,

$$Twin_1 = \left(\frac{\sum_{i \in (P_{2it-1}=1)} \Pr(P_{1it}=1 | P_{2it-1}=1)}{\sum_i P_{2it-1}} \right) - \left(\frac{\sum_{i \in (P_{2it-1}=0)} \Pr(P_{1it}=1 | P_{2it-1}=0)}{\sum_i (1-P_{2it-1})} \right)$$

The first term in $Twin_1$ is the average probability of the first twin being poor conditional on that the second twin was poor the previous year. The second term is the average probability of poverty for the first twin if the second twin was not identified as poor. Finally, the difference between the averages among the respective subgroups is calculated. At this point it is important to note that it is the averages among two different *groups* that are compared. The method is certainly different from, and should not be confused with, how unobserved heterogeneity is taken into account in other twin based methods. In studies dealing

with ability bias in estimates for the return to schooling, differences between twins are used to control for unobserved heterogeneity. Instead of differencing away heterogeneity, $Twin_1$ uses the similarity among pairs of twins to identify the family specific heterogeneity in poverty persistence, i.e. it measures family specific heterogeneity.

The reason that $Twin_1$ identifies family specific heterogeneity is that it is not possible to have state dependence between the twins under the Twin-State Independence Assumption. Common traits, characteristics or innate abilities can, however, have influenced the risk of poverty for both the twins. If one's background matters for the risk of poverty, it is expected that the probability is higher for those twins who have a twin sibling that was poor the previous year. Accordingly, it is expected that $Twin_1$ will be positive.

As described earlier, the persistence in poverty is explained by true state dependence and heterogeneity. In this study, the persistence in poverty is estimated using ASD , and the part that may be due to family specific heterogeneity is estimated by means of $Twin_1$.

$$ASD_1 = TSD_1 + Twin_1 \quad (10)$$

If family specific heterogeneity did not matter at all, there would be no difference between the averages in $Twin_1$. Accordingly, the measure of $Twin_1$ would be zero, and TSD_1 would be the explanation for state dependence. If family specific heterogeneity were to explain almost all of the persistence in poverty, this measure would tend to approach the estimate of ASD_1 , as measured above. TSD_1 would, as a consequence, be very small. Accordingly, with the extra information on twins, it is

possible to distinguish family specific heterogeneity from true state dependence. It is also possible to compare the results with estimates from methods used in the previous literature. *ASD* and *GSD* are here calculated and compared with the measure of heterogeneity that can be calculated from the twin information.

3. Data

The data used for this study is a combination of administrative data covering the total Swedish population and survey data. The included individuals are twins born between 1949 and 1958, found in the Swedish Twin Registry. Unfortunately the Swedish Twin Registry does not include information on whether the twins are monozygotic or dizygotic for individuals born after 1958. Dizygotic twins are genetically not more alike than ordinary siblings, while monozygotic twins are identical with respect to their genes. Since this distinction is potentially important for our purposes, individuals born after 1958 are not included in this study. Focus is, of course, on the monozygotic sample, but results are also presented for the measures of true state dependence and heterogeneity for the dizygotic sample. The reason is to see whether the potential extra homogeneity is important.¹² The population is born exclusively in Sweden, which makes it dubious to generalize results to immigrants.¹³

¹²It is not necessarily the case that monozygotic twins are more homogenous, even though the genes are more alike. Psychological reasons could, for example, create a larger need for monozygotic twins to diverge in decisions and life style to underline that they are in fact different individuals and not identical.

¹³Data on immigrants can be included through spouses. There are, however, relatively few cases to rely on, and also no couples where both are immigrants would be included. In this study, no attempts are made to say anything about poverty among immigrants. See Hansen & Löfström (2003), for a study on immigrants' welfare participation in Sweden.

Information concerning, for example, income, social assistance and education from 1994 until 1999 are attached to the population of twins. This information is based on different registers for the total Swedish population and also included in the longitudinal database, Louise.¹⁴ Biological parents are connected to the twins through "Several Generations Registry".¹⁵ Data for the years 1960 and 1970 are also included. These data come from a nationwide census called "Population and Housing Census".¹⁶ Information from 1994 until 1999 is also included for spouses. The data were linked and matched by Statistics Sweden. All the data, except the twin information, also come from Statistics Sweden.

One problem with empirical investigations of poverty is that it is necessary to find a measure that captures a definition of poverty. This can, quite easily, become rather complex. It is, for example, possible to define a measure of poverty in absolute terms or in relative terms, where poverty also depends on a relative position in society. Further, it is not obvious that the measure should be based on financial resources, since these are not necessarily a guarantee for a rich life. Even though there are no official measures of poverty, a commonly used poverty line is 50 percent of the median household disposable income measured per consumption unit.¹⁷ It is common to weigh the incomes with some kind of equivalence scale since larger families need more resources to achieve a reasonable living standard. The consequences of using different equivalence scales are, for example, analyzed in Coulter *et al.*

¹⁴The database is described (in Swedish) in "Bakgrundsfakta till arbetsmarknads- och utbildningsstatistiken, 2002:2. Statistics Sweden."

¹⁵In Swedish, "Flergenerationsregistret".

¹⁶In Swedish, "Folk- och Bostadsräkningen, FoB".

¹⁷There is substantial literature on different poverty measures. Bennet (2001), chapter four in "Social Rapport 2001", published by "Socialstyrelsen", The National Board of Health and Welfare in Sweden, includes a deeper discussion of different measures of poverty in the Swedish context. This study is, however, limited to measures that can be defined from administrative data.

(1992) and Phipps (1993). Not surprisingly the measure of poverty depends on how the equivalence scale is constructed.

The analysis in this study is performed using two different measures of poverty. Poverty is defined either as based on whether the individual received social assistance during the year or, whether the individual had a disposable income below 60 percent of the median of the sample. For the measure based on disposable income, an equivalence scale is used. This is based on norms defined by "Socialstyrelsen", The National Board of Health and Welfare in Sweden. Further, to define poverty, the disposable incomes of the members of the family are added together. Then, the sum of the disposable incomes is multiplied by the individual's consumption weight and divided with the sum of the consumption weights for the family.¹⁸ An individual is identified as poor if he/she does not reach 60 percent of the median consumption weighted disposable income for the sample.¹⁹ There are three reasons for the choice of 60 percent instead of 50 percent. The first reason is to make the results more comparable with Cappellari & Jenkins (2004), since they used 60 percent as the poverty line. A second reason is Cappellari & Jenkins' (2002) finding that to use a low line could hide the importance of heterogeneity in the measures.²⁰ Finally, it would be difficult for the estimations to converge if too few individuals are actually observed as poor. A drawback with this poverty measure, and

¹⁸The consumption weights are based on norms defined by "Socialstyrelsen", The National Board of Health and Welfare in Sweden. If the family only consists of one adult, the weight is 1.16. For two or more adults each adult is weighted 0.96. Children, 0-3, 4-10, 11-17 years old add respectively 0.56, 0.66 and 0.76.

¹⁹With this poverty line, exit or entry into poverty could be identified for the individual, even though the economic situation has almost not changed at all. Jenkins (2000) suggests a method to avoid threshold effects due to arbitrarily defined poverty lines. The idea is to only define an exit if the disposable income reaches 10% above the poverty line, and define entry only if disposable income does not reach 90% of the poverty line. This method is not applied here since the main focus is poverty persistence, and introducing this idea would make the results less transparent.

²⁰The suggested reason for this is that heterogeneity could be less different among individuals above and below the poverty line when it is set very low.

the data, is that cohabiting partners are only identified if they are married or have at least one common child. This means that some families could have an extra member with an important income.

As mentioned previously, the second measure of poverty used in this study is whether the individual was a receipt of social assistance. This is the last resort of public assistance when the other systems, such as unemployment benefits, are not enough or not applicable. To get social assistance in Sweden, the individual is required to be trying to support him/herself as far as he/she is able. Usually savings have to be used before social assistance is granted. Further, social assistance is not granted if the individual has a family, such as a cohabiting partner, that can assist. To receive social assistance, the individual has to apply for it. This means that it is possible that individuals who are entitled to receive assistance do not apply for it, which is, of course, a drawback when using the presence of social assistance as a measure of poverty.

As social assistance is not granted to those who have a cohabiting partner who can provide financial support, the measure based on social assistance does not have the problem mentioned for the measure constructed from disposable income. In other words, it avoids the problem that an extra family member can be present with an unknown income. The preferred measure is the one indicating that the individual received social assistance, as it is based on the information available to the local social welfare worker concerning the need for social assistance.

The data analyzed include 1749 monozygotic pairs of twins, and 2620 dizygotic pairs of twins with the same sex, born between 1949 and 1958.²¹ Information for the parents is included for 1960 and 1970. The

²¹Dizygotic pairs of twins that are not of the same sex are not included in the study. The reason is that the results from pairs of dizygotic twins will be compared to results

twins were, accordingly, between 2 and 11 years in 1960, and between 12 and 21 years old in 1970. Where one of the twins was self-employed, the observations are excluded. In such a case, the twin has a different control over his/her yearly income than employees usually have. In addition, only pairs where both twins were alive at least until year 2000 are included in the analysis. With these restrictions, 912 observations of poverty, using the measure based on social assistance, are identified for the monozygotic twins for the period 1994 - 1999. For the same period 1056 observations are included of a monozygotic twin that had a disposable income lower than 60 percent of the median. Note that these numbers are added for both the twins and their twin siblings. The numbers are also added for the period 1994 - 1999 and, in the case of poverty persistence, fewer pairs of twins have experienced poverty. Descriptive statistics for the first set of monozygotic twins can be found in *Table 1*.

from pairs of monozygotic twins and adding this difference would make it less relevant to compare the estimates.

Table 1. Descriptive statistics.

Monozygotic sample		
Number of observations: 6491		Twin 1
Variabels	Mean	Standard deviation
<i>receiving social assistance (1=yes)</i>	0.0481	0.2139
<i>Below 60 % of median disposable income (1=yes)</i>	0.0530	0.2240
Region (1=Stockholm, Malmö or Gothenburg)	0.4605	0.4985
Education (1=upper secondary school, less than 3 years)	0.4765	0.4995
Education (1=post secondary school and post graduate education)	0.3310	0.4706
Married (1=married)	0.5978	0.4904
Number of children 0-3 years	0.0739	0.3047
Number of children 4-6 years	0.1246	0.3653
Number of children 7-10 years	0.2770	0.5321
Number of children 11-15 years	0.4623	0.6677
Number of children 16-17 years	0.1725	0.3895
Months unemployed (i.e. days unemployed during year/30)	1.0254	2.9783
Partner not identified (1=cohabiting partner not identified)	0.3243	0.4681
Education of partner (1=upper secondary school, less than three years)	0.4717	0.4993
Education of partner (1=post secondary school and post graduate education)	0.3393	0.4735
Months of unemployed for partner (i.e. days unemployed during year/30)	0.6526	2.3791
Age	43.9111	3.0796
Age of partner	44.0337	5.7289
Female (1=female)	0.5542	0.4971
Mother not in the labor market 1960	0.8718	0.3343
Mother not in the labor market 1970	0.4200	0.4936
Education of mother, 1970 (1= upper secondary school, less than 3 years)	0.1524	0.3595
Education of mother, 1970 (1= post secondary school and post graduate education)	0.0545	0.2271
Education of father, 1970 (1= upper secondary school, less than 3 years)	0.2513	0.4338
Education of father, 1970 (1= post secondary school and post graduate education)	0.0706	0.2562

Using poverty measured as receiving social assistance, a little less than 5 percent of the observations are identified as poor. When poverty measured as a disposable income less than 60 percent of the median income is used, slightly more than 5 percent of the observations are identified as situations of poverty.

The different measures of poverty include, to a large extent, different individuals. Out of those who received social assistance, only 15.1 percent were also counted as poor according to the measure based on disposable income. In the same way, only 13.7 percent of those with a disposable income below 60 percent of the median actually received social assistance during the year.

4. Results

The most important result in this study is that family specific heterogeneity is estimated to be the reason for between 24 - 31 percent of the poverty persistence, when social assistance is used as a measure of poverty. Before explaining these measures in detail, the presentation of the results includes both parameters of characteristics that are assumed to affect poverty transitions and the correlation of error terms between the equations. Even though the focus is on a sample of monozygotic twins, estimations are also made for dizygotic twins of the same sex. This section also includes estimates for two different poverty measures. Results of the four variate probit model, for the monozygotic twin sample, can be found in *Table 2*.

Table 2. Estimates.

Monozygotic sample <i>t</i> = 1996 & 1998	<i>Poverty measured as received social assistance during the year</i>			<i>Poverty measured as income less than 60% of median disposable income</i>		
	$P_{1t-1} = 1$	$P_{1t} = 1$ given $P_{1t-1} = 1$	$P_{1t} = 1$ given $P_{1t-1} = 0$	$P_{2t-1} = 1$	$P_{2t} = 1$ given $P_{2t-1} = 1$	$P_{2t} = 1$ given $P_{2t-1} = 0$
Variables, (measured at <i>t</i> -1) ^a	$(\beta_1 = \beta_2)^b$	$(\gamma_{11} = \gamma_{21})^b$	$(\gamma_{12} = \gamma_{22})^b$	$(\beta_1 = \beta_2)^b$	$(\gamma_{11} = \gamma_{21})^b$	$(\gamma_{12} = \gamma_{22})^b$
Region (1=Stockholm, Malmö or Gothenburg)	-0.0191 -0.28	0.0334 0.21	-0.0576 -0.68	-0.0507 -0.71	-0.2638 -1.56	-0.0820 -0.99
Education (1=upper secondary school, less than 3 years)	-0.3234*** -3.91	0.2540 1.39	-0.1397 -1.41	-0.0246 -0.28	-0.0490 -0.26	0.0417 0.40
Education (1=post secondary school and post graduate education)	-0.6711*** -6.13	-0.3205 -0.98	-0.3832*** -3.08	-0.3237*** -2.75	0.0994 0.36	-0.1325 -1.02
Married (1=married)	-0.0166 -0.15	0.0764 0.29	0.0898 0.58	0.2631** 2.34	-0.0163 -0.06	0.2559** 1.97
Number of children 0-3 years	0.2583*** 2.71	0.4285* 1.76	0.1346 1.07	0.6201*** 7.91	0.0092 0.04	0.4326*** 4.31
Number of children 4-6 years	-0.0795 -0.85	0.2960 1.35	0.0929 0.99	0.3909*** 5.98	-0.0009 -0.01	0.4427*** 5.30
Number of children 7-10 years	0.0983* 1.63	0.0521 0.31	0.0168 0.21	0.3172*** 5.76	0.1809 1.30	0.2890*** 4.11
Number of children 11-15 years	0.0336 0.63	0.0655 0.53	0.2469*** 4.27	0.3853*** 8.78	-0.0658 -0.51	0.2007*** 3.43
Number of children 16-17 years	0.1685** 2.11	-0.0609 -0.35	0.0515 0.52	0.2574*** 3.75	0.0480 0.29	0.3065*** 2.99
Months unemployed during year (i.e. days/30)	0.0680*** 9.10	0.0244 1.04	0.0350*** 3.30	0.0383*** 4.34	-0.0122 -0.55	0.0100 0.84
Partner not identified (1=cohabiting partner not identified)	-1.6994 -1.14	-1.3467 -0.49	-0.2591 -0.13	-0.8271 -0.61	-1.5887 -0.57	-2.2675 -1.62
Education of partner, (1=upper secondary school, less than three years)	-0.1131 -1.07	-0.0534 -0.22	-0.2432* -1.95	-0.3606*** -3.98	-0.2087 -1.03	-0.0642 -0.56
Education of partner, (1=post secondary school and post graduate education)	-0.2494* -1.83	-0.1107 -0.35	-0.3581** -2.44	-0.7020*** -5.55	0.0384 0.11	-0.4742*** -3.25
Partner unemployed (1=one or more days of unemployment during year)	0.0523*** 4.37	-0.0301 -0.88	0.0571*** 3.86	0.0295** 2.40	0.0186 0.65	0.0051 0.34

Table continues on the following page.

Table 2. Continues...

Variables	Poverty measured as received social Assistance during the year			Poverty measured as income less than 60% of median disposable income		
	$P_{1t-1} = 1$ $(\beta_1 = \beta_2)^b$	$P_{1t} = 1$ given $P_{1t-1} = 1$ $(\gamma_{11} = \gamma_{21})^b$	$P_{1t} = 1$ given $P_{1t-1} = 0$ $(\gamma_{12} = \gamma_{22})^b$	$P_{2t-1} = 1$ $(\beta_1 = \beta_2)^b$	$P_{2t} = 1$ given $P_{2t-1} = 1$ $(\gamma_{11} = \gamma_{21})^b$	$P_{2t} = 1$ given $P_{2t-1} = 0$ $(\gamma_{12} = \gamma_{22})^b$
Age	0.2114 0.72	-0.9172 -1.24	0.2861 0.78	0.3403 1.15	-0.6146 -0.81	-0.1174 -0.32
Age – squared	-0.0023 -0.67	0.0105 1.19	-0.0030 -0.70	-0.0041 -1.18	0.0071 0.79	0.0012 0.29
Age of partner (0 if no partner present)	-0.0967 -1.39	-0.1349 -1.01	-0.0235 -0.25	-0.0602 -1.00	-0.0817 -0.59	-0.1222** -1.96
Age of partner – squared	0.0011 1.36	0.0022 1.36	0.0002 0.24	0.0009 1.32	0.0013 0.78	0.0015** 2.08
Female (1=female)	0.1368* 1.73	-0.3150* -1.74	0.1143 1.23	0.0570 0.68	-0.1884 -0.96	0.0940 0.96
Mother not in the labor market 1960	0.0511 0.45			-0.0132 -0.11		
Mother not in the labor market 1970	-0.1150 -1.48			0.1700** 2.11		
Education of mother, 1970 (1= upper secondary school, less than 3 years)	0.2120** 2.14			0.1885* 1.86		
Education of mother, 1970 (1= post secondary school and post graduate education)	0.1241 0.59			0.3466** 2.44		
Education of father, 1970 (1= upper secondary school, less than 3 years)	0.0753 0.88			-0.1758** -1.96		
Education of father, 1970 (1= post secondary school and post graduate education)	-0.1504 -0.85			-0.1803 -1.29		
Constant	-4.4809 -0.73	22.2944 1.37	-8.3490 -1.04	-8.2946 -1.34	15.1985 0.95	2.5894 0.34
Number of observations	3238			3238		
Log likelihood	-1846.89			-1824.66		

Note: Estimated coefficients are found in the first line for each variable. The second row includes the asymptotic t -ratios. The standard errors are corrected for repeated observations from the same twins over the years. Coefficients that are significantly different from zero at 10, 5 and 1%-levels are marked with *, ** and ***. Likelihood ratio tests confirm that the instruments can be excluded from the transition equations for both measures of poverty. The test statistics were estimated to 14.64 and 17.63, which both are less than the critical value of 21.03 in the χ^2 -distribution for the significance level of 5% with 12 degrees of freedom. a) Variables concerning unemployment are measured at $t-2$. b) Parameters for variables are constrained to be the same for the initial condition equations for each set of twins, i.e. $\beta_1 = \beta_2$. The same concerns the parameters for the transition equations for the two groups of twins, i.e. $\gamma_{11} = \gamma_{21}$ and $\gamma_{12} = \gamma_{22}$.

Table 2 includes estimates for both measures of poverty. The results are based on estimates where the years 1996 and 1998 are used as transition years with 1995 and 1997 as the years for the initial condition.²² Note that in the estimation β_1 , are constrained to be equal to β_2 , i.e. the coefficients in equations (1) and (2) are assumed to be the same. In the same way γ_{11} and γ_{12} are constrained to be equal to γ_{21} respective γ_{22} for equations (3) and (4). This can be done since there are no reasons to expect a systematic difference between the parameters for the different groups of twins.

As indicated in the description of the model, the explanatory variables included are lagged one period in all the equations. Variables concerning the unemployment of the individual and his/hers spouse are, however, an exception. It could, for example, be possible that the individual enters unemployment and poverty simultaneously due to some unobserved event. Accordingly, the variables concerning unemployment are lagged two steps to reduce the risk that equation (1) and (2) include potentially endogenous variables.

The overview of the results indicates that many parameters are not significantly different from zero. For the transition equations, this is something that also has been found in previous literature (Stewart & Swaffield, 1999 and Cappellari & Jenkins, 2004). The model allows different estimates of the parameters depending on the poverty status in previous years. With very few observations of poverty the previous year, the estimation of the parameters for some of the dummy variables has to rely on very few observations. This could be an explanation of the difficulty to obtain significant coefficients in the estimates for the transition out of poverty.

²²The data also permit estimates using 1997 and 1999 as years for transition and 1996 and 1998 as the base year. These estimates, as well as estimates for the dizygotic twins, can be requested from the author.

Focusing on the results when poverty is measured by social assistance, not surprisingly, higher education reduces the risk of entering into poverty. The coefficient is significantly different from zero at the 1 percent significance level using a double sided test.²³ In this sample, however, higher education does not have a significant effect on the transition out of poverty. The experience of being unemployed two years previously increases the risk of being poor for individuals who were not observed as poor the previous year. This is also the case for a cohabiting spouse who is unemployed. Both coefficients are significantly different from zero at the 1 percent probability level. For the probability of leaving the state of poverty, having a spouse who was unemployed the previous year does not seem to matter.²⁴

When poverty is measured using disposable income, the coefficients of many variables are, again, not significantly different from zero. However, the risk of entering into poverty seems higher when children are present. All the parameters for children are significantly different from zero at the 1 percent level. When the individual is already poor, children do not seem to affect the probability that he/she will remain in that state.²⁵

The correlations of the error terms for the equations are included in *Table 3*.

²³All the subsequently mentioned tests are double sided tests unless something else is explicitly written.

²⁴A Wald test indicates a significant difference between the coefficients depending on the poverty status the previous year. The Wald statistic is estimated to 5.88 compared with the critical value of 3.84 in the χ^2 -distribution for the 5 percent level.

²⁵Wald tests for the number of children between 4 and 6 years old, and 11 and 15 years old indicate significant differences between the coefficients depending on the poverty status the previous year. The Wald statistics are estimated to 5.45 and 4.05 compared with the critical value of 3.84 in the χ^2 -distribution for the 5 percent level.

Table 3. Correlations of error terms

Monozygotic sample <i>t</i> = 1996 & 1998		<i>Poverty measured as received social assistance during the year</i>		<i>Poverty measured as income income less than 60% of median disposable income</i>	
		Estimate	t	Estimate	t
twin 1 poor t-1, twin 2 poor t-1	ρ_{21}	0.3160***	4.33	0.3504***	4.93
twin 1 poor t-1, twin 1 poor t	ρ_{31}	-0.1022	-0.36	-0.0530	-0.24
twin 1 poor t, twin 2 poor t-1	ρ_{32}	0.0591	0.63	0.1961**	2.15
twin 1 poor t-1, twin 2 poor t	ρ_{41}	0.2200**	2.54	0.1205	1.23
twin 2 poor t-1, twin 2 poor t	ρ_{42}	-0.1473	-0.46	-0.0951	-0.36
twin 1 poor t, twin 2 poor t	ρ_{43}	0.3258***	3.58	0.2789***	2.85

The correlations of the error terms between the twins are, in all but one case, significantly different from zero. This is the case, regardless of the type of poverty measure used. The signs of the correlations of the error terms are positive, which indicates a positive correlation of unobservables between the twin siblings. The correlation of the error terms for the initial condition equations between the twins, are significantly different from zero at the 1 percent probability level. This is also the case for the transition equations between the twins. The correlation of the error terms between equations one and three, and two and four, are not significantly different from zero.

Table 4. Measures of state dependence and heterogeneity

Monozygotic sample	Poverty measured as received social assistance during the year		Poverty measured as income less than 60% of median disposable income	
	Twin 1	Twin 2	Twin 1	Twin 2
ASD _{96, 98}	0.5908	0.5610	0.6503	0.6349
GSD _{96, 98}	0.5019	0.4729	0.6267	0.5930
Difference (ASD-GSD)	0.0889	0.0881	0.0236	0.0419
Share of ASD due to heterogeneity	15.05%	15.70%	3.63%	6.60%
Twin _{96, 98}	0.1613	0.1726	0.1424	0.1553
Share of ASD due to heterogeneity	27.30%	30.77%	21.90%	24.46%
ASD _{97, 99}	0.5257	0.5230	0.6576	0.6548
GSD _{97, 99}	0.4663	0.4532	0.6080	0.5777
Difference (ASD-GSD)	0.0594	0.0698	0.0496	0.0771
Share of ASD due to heterogeneity	11.30%	13.35%	7.54%	11.77%
Twin _{97, 99}	0.1238	0.1244	0.1496	0.1520
Share of ASD due to heterogeneity	23.55%	23.79%	22.75%	23.21%

Note: index 96, 98 refers to measures based on estimation with transition equation 1996 and 1998 and year 1995 and 1997 as initial condition equations. Index 97, 99 is based on a model with 1997 and 1999 as transition equations.

The estimates for state dependence and heterogeneity for the monozygotic twins are included in Table 4.²⁶ Estimates for the measures are included both when 1996 and 1998, and when 1997 and 1999 are used as transition equations. The measure for the overall state dependence, *ASD*, is estimated to about 0.52 – 0.59 when social assistance is used as the poverty measure. Accordingly, an individual who experienced poverty during the preceding year has a substantially higher risk of staying in poverty than an individual who was not poor the previous year has of entering poverty. The estimates of genuine state dependence, *GSD*, are about 0.06 – 0.09 lower. The estimates for the family specific heterogeneity, estimated using the twin method (as in *Twin*₁), are about

²⁶In the calculations, some observations were dropped since the denominator was approaching zero. If the denominator was less than 0.001, the observation was dropped from the calculations.

0.12 – 0.17. In other words, the risk of being poor is about 0.12 – 0.17 higher if the twin sibling was observed to be poor the previous year compared with if he/she was not observed as poor. By using equation (10), it follows that 24 – 31 percent of the aggregate state dependence is due to family specific heterogeneity.²⁷ The *GSD* measure indicates that 11 – 16 percent of the poverty persistence is due to heterogeneity.²⁸

When the poverty measure is based on disposable income, the difference between *ASD* and *GSD*, for the monozygotic sample, is 0.02 – 0.08. The twin based measure of heterogeneity gives a result of 0.14 – 0.16. Accordingly, the twin method attaches 22 – 24 percent of the poverty persistence as due to family specific heterogeneity.²⁹ Whereas the *GSD* measure indicates that only 4 – 12 percent is due to heterogeneity.³⁰

The same measures are also calculated for the dizygotic sample. The results are included in *Table 5*.

²⁷The shares are calculated, from *Table 4*, according to, $0.1244/0.5230 \approx 24\%$ and $0.1726/0.5610 \approx 31\%$.

²⁸From *Table 4*, $0.0594/0.5257 \approx 11\%$ and $0.0881/0.5610 \approx 16\%$.

²⁹From *Table 4*, $0.1424/0.6503 \approx 22\%$ and $0.1553/0.6349 \approx 24\%$.

³⁰From *Table 4*, $0.0236/0.6503 \approx 4\%$ and $0.0771/0.6548 \approx 12\%$.

Table 5. Measures of state dependence and heterogeneity

Dizygotic sample	Poverty measured as received social assistance during the year		Poverty measured as income less than 60% of median disposable income	
	Twin 1	Twin 2	Twin 1	Twin 2
ASD _{96, 98}	0.5288	0.5070	0.6738	0.6553
GSD _{96, 98}	0.4239	0.3749	0.5832	0.5342
Difference (ASD-GSD)	0.1049	0.1321	0.0906	0.1211
Share of ASD due to heterogeneity	19.84%	26.06%	13.45%	18.48%
Twin _{96, 98}	0.0847	0.0819	0.0627	0.0973
Share of ASD due to heterogeneity	16.02%	16.15%	9.31%	14.85%
ASD _{97, 99}	0.5723	0.4436	0.6665	0.6649
GSD _{97, 99}	0.4828	0.3113	0.5713	0.5578
Difference (ASD-GSD)	0.0895	0.1323	0.0951	0.1071
Share of ASD due to heterogeneity	15.63%	29.82%	14.27%	16.11%
Twin _{97, 99}	0.0738	0.1068	0.0920	0.1168
Share of ASD due to heterogeneity	12.90%	24.08%	13.8%	17.57%

Note: index 96, 98 refers to measures based on estimation with transition equation 1996 and 1998 and year 1995 and 1997 as initial condition equations. Index 97, 99 is based on a model with 1997 and 1999 as transition equations.

When social assistance is used as the measure for poverty, *ASD* is estimated to be 0.44 – 0.57, while *GSD* is 0.31 – 0.48 for the dizygotic sample. The difference between the measures is about 0.09 – 0.13. This indicates that about 16 – 30 percent of the persistence of poverty is due to heterogeneity according to the method previously used in the literature.³¹ The twin method attaches about 0.07 – 0.11 as due to family specific heterogeneity, which suggests 13 – 24 percent of the poverty persistence is due to this type of heterogeneity.³²

When disposable income is used to construct the poverty measure, the difference between *ASD* and *GSD* is about 0.09 – 0.12. Accordingly, about 13 – 18 percent of the poverty persistence is due to heterogeneity.³³

³¹From Table 5, $0.0895/0.5723 \approx 16\%$ and $0.1323/0.4436 \approx 30\%$.

³²From Table 5, $0.0738/0.5723 \approx 13\%$ and $0.1068/0.4436 \approx 24\%$.

³³From Table 5, $0.0906/0.6738 \approx 13\%$ and $0.1211/0.6553 \approx 18\%$.

The measure for family specific heterogeneity based on the twin method is 0.06 – 0.12. This indicates that about 9 – 18 percent of the poverty persistence is due to this type of heterogeneity.³⁴

The twin method ascribes a smaller amount of the causes of the persistence of poverty to family specific heterogeneity when dizygotic twins are used instead of monozygotic twins. Dizygotic twins are, however, not genetically more alike than ordinary siblings. Accordingly, a substantial part of the innate abilities are not included in the measure.

When using the measure of poverty based on disposable income, the amount of poverty persistence due to heterogeneity is estimated to be slightly lower than when social assistance is used. The results underline the importance of true state dependence in poverty, irrespective of the poverty measure used. This conclusion applies regardless of whether the twin method is used or not.

Biewen (2004) estimates, for a German data set, that the probability of being poor if the individual was already poor the previous year, is about 45 percent higher than for individuals who were not poor previously. He estimates a joint dynamic model of poverty, employment status and household composition. Observed and unobserved heterogeneity are estimated to be the reason for about half of the persistence of poverty. For British data, Cappellari & Jenkins (2004) estimate *ASD* to 0.526 and *GSD* to 0.310 when using a poverty line set to 60 percent of median income.³⁵ Accordingly, about 41 percent of the poverty persistency is estimated to be due to heterogeneity.

³⁴From *Table 5*, $0.0627/0.6738 \approx 9\%$ and $0.1168/0.6649 \approx 18\%$.

³⁵They use the McClements equivalence scale, and post-tax and post-transfers income.

In this study, when poverty is measured as receiving social assistance, the twin method assigns about 24 – 31 percent of the persistence of poverty to family specific heterogeneity. When poverty was defined as a disposable income lower than 60 percent of the median disposable income, the amount due to heterogeneity was estimated to be even lower, i.e. 22 – 24 percent. Thus, true state dependence is relatively more important than family specific heterogeneity in explaining poverty persistence for the sample of Swedish twins used here.

Cappellari & Jenkins (2002) tested a number of different poverty lines and found that, for the lowest poverty threshold line, "*GSD* was estimated to be even larger than *ASD* (albeit only slightly)". They suggest that heterogeneity would vary less among individuals below and above the poverty line when it was set very low. The results in this study suggest that heterogeneity could be underestimated when using the method applied by Cappellari & Jenkins (2004) to distinguish between true state dependence and heterogeneity. The reason is that the twin method identifies a lower bound for individual specific heterogeneity and still the measure is estimated to be higher compared with the estimate based on the method in Cappellari & Jenkins (2004). Another possible reason for the rather low estimates for heterogeneity, when their method is used for the monozygotic sample, could be that it is a quite small sample. The dizygotic sample seems to produce higher estimates for heterogeneity when Cappellari & Jenkins' (2004) method is used. It is possible that the twin method is less sensitive to a small sample size.

It is also worth taking note of the difference between using monozygotic and dizygotic twins in the twin method. Not surprisingly, the heterogeneity, using information on dizygotic twins, is estimated to be less compared with the monozygotic sample. The monozygotic twins are more homogenous and accordingly more family specific heterogeneity is identified. When estimating the return to education, using data on Swedish twins, Isacsson (2004) also concludes that the information about zygosity seems to be important.

The twin method relies on an assumption that the experience of one twin does not, in itself, increase the probability for the sibling twin to experience poverty the following year. When it comes to social assistance, a number of questions could be raised concerning the independence of receiving social assistance for twin siblings. It could, for example, be the case that one sibling twin opens the other twin's eyes to the possibility to get assistance. It could also be that to be in receipt of social assistance is accompanied by some degree of embarrassment which is partly taken away by the fact that the first twin sibling is already receiving assistance. If this were the case, the Twin-State Independence Assumption would not be fulfilled. That is, *the experience* of poverty for the sibling twin would increase the probability for poverty for the other twin. On the contrary, the method is based on the assumption that it is common characteristics that increase the probability of poverty for the sibling twin and not the experience in itself. Comparing the estimates from the monozygotic and the dizygotic samples does, however, indicate that there is a substantial difference in family specific heterogeneity that hardly can be explained by a dependency between receiving social assistance for twin siblings. There are no obvious reasons why the possible dependence in receiving social assistance between the siblings should be higher for monozygotic twins.

5. Concluding remarks

This study focuses on the persistence of poverty in Sweden. The purpose is to distinguish between two different reasons why individuals who are found to be poor one year are more likely to continue to be poor the following year. One suggested reason for poverty persistence is that individuals with certain characteristics are more likely always to be poor, as these characteristics hardly change. In this case, the reason for continuing poverty would be heterogeneity. Another reason for poverty persistence is true state dependence, i.e. that the experience of poverty, in itself, causes a higher risk of remaining in poverty in the coming years. Distinguishing between these two reasons is important for designing an effective policy to handle poverty.

This study uses a sample of twins to distinguish between true state dependence and family specific heterogeneity. It exploits the fact that monozygotic twins have very similar backgrounds and also are genetically the same. The similarity between the twins is used to identify the part of poverty persistence that is due to family specific heterogeneity.

Using social assistance as an indicator of poverty, the probability of remaining poor is estimated to be 0.52 – 0.59 higher than the probability of becoming poor when a state of poverty was not experienced the previous year. This higher risk is likely to be due to both heterogeneity and true state dependence. The risk of poverty is estimated to be about 0.12 – 0.17 higher if the monozygotic sibling twin experience poverty in the previous year than if he/she did not. Since state dependence cannot explain a higher risk, the latter result must be due to family specific heterogeneity. Accordingly, about 24 – 31 percent of the poverty persistence is caused by this type of heterogeneity. When the measure of poverty is based on

disposable income, the twin method assigns 22 – 24 percent of the poverty persistence to family specific heterogeneity.

The results can be compared with Biewen (2004) who finds, for a German data set, that heterogeneity explains half of the poverty persistence in the sample. Cappellari & Jenkins (2004) explain 41 percent of the persistence in poverty by heterogeneity. Even if it is difficult to have an opinion on whether the results in this study indicate little, or a lot, of the poverty persistence that is due to family specific heterogeneity, it has to be remembered that the twin method only can identify *family* specific heterogeneity. The acquired experiences that differ between the twins, and which also can affect the risk of poverty, are not captured in the measure. Accordingly, the twin method identifies the part of poverty persistence that is due to family specific heterogeneity as well as, a lower bound for individual specific heterogeneity. For the monozygotic sample, the method used in Cappellari & Jenkins (2004) attaches only 4 – 12 percent to individual specific heterogeneity when disposable income is used as measure of poverty. As long as the Twin-State Independence Assumption is valid, individual specific heterogeneity seems to be more important than that. The conclusion concerning the importance of true state dependence seems, however, to be very solid.

One concern about the estimates is the limited sample of twins used both with regard to age and to the lack of immigrants. It is possible that the persistence of poverty differs for younger individuals and immigrants. In the same way, it is possible that the relative importance of family specific heterogeneity and true state dependence could differ. The twin method also relies on it being possible to generalize the results to the overall population. Growing up with a sibling twin is certainly not the same as growing up as an only child. It is, however, not necessarily the case that

these differences would affect the poverty persistence, or true state dependence.

It would be interesting to see estimates for twin samples covering younger individuals and immigrants. It would also be interesting to repeat the study for different countries with different welfare systems. It is possible that heterogeneity plays a different role in a country where there is a smaller public sector. Another area for future research is to investigate the reasons for true state dependence. Knowing that true state dependence is important encourage a policy that avoids people entering into poverty. Knowing *why* true state dependence occurs would, of course, also help in designing a system that reduces the persistency of poverty.

6. References

Bakgrundsfakta till arbetsmarknads- och utbildningsstatistiken, 2002:2. Statistics Sweden.

Bennet, E. W. (ed.), (2001), *Social Rapport 2001*, Socialstyrelsen, Linköping: Linköpings Tryckeri AB.

Biewen, M. (2004), Measuring State Dependence in Individual Poverty Status: Are There Feedback Effects to Employment Decisions and Household Composition?, *IZA Discussion Paper Series*, No. 1138. Institute for the Study of Labor, Bonn.

Cappellari, L. & S. P. Jenkins, (2002), Modelling Low Income Transitions, *ISER Working Papers 2002-8*, Institute for Social and Economic Research, University of Essex.

Cappellari, L. & S. P. Jenkins, (2003), Multivariate probit regression using simulated maximum likelihood, *The Stata Journal*, 3(3), 278-294.

Cappellari, L. & S. P. Jenkins, (2004), Modelling Low Income Transitions, *Journal of Applied Econometrics*, 19: 593-610.

Coulter, F. A. E., F. A. Cowell & S. P. Jenkins, (1992), Equivalence scale Relativities and the Extent of Inequality and Poverty, *The Economic Journal*, 102, 1067-1082.

Hansen, J. & M. Löfström, (2003), Immigrant Assimilation and Welfare Participation. Do Immigrants Assimilate Into or Out of Welfare?, *Journal of Human Resources*, v. 38, issue 1, 74-98.

Isacsson, G. (2004), Estimating the Economic Returns to Educational Levels Using Data on Twins, *Journal of Applied Econometrics*, 19: 99-119.

Jenkins, S. P. (2000), Modelling Household Income Dynamics, *Journal of Population Economics*, 13: 529-567.

Phipps, S. A. (1993), Measuring Poverty Among Canadian Households. Sensitivity to Choice of Measure and Scale, *Journal of Human Resources*, Winter, v. 28, issue 1, 162-184.

Stewart, M. B. & J. K. Swaffield, (1999), Low Pay Dynamics and Transition Probabilities, *Economica*, 66, 23-42.